

Z. Karvalics László

A Nagy Adat-jelenség társadalomtudományi lehorgonyzásához

Hogy itt mi lesz, ki érti ezt?

Gondolat helyett gondolatmenet...

Ki érti ezt? Hogy itt mi lesz ...?

(Bereményi Géza – Cseh Tamás: F. M. Dosztojevszkij és az ördög)

A „Nagy Adat” (Big Data) fogalma nem egy új tudományos diskurzus részeként született meg, mint például a TMI (*Too Much Information*) számos sebtől vérző kvázielméletcskéje: az adattermelés és -feldolgozás új mennyiségi dimenzióit megnyitó alaprendszerek különöségét megragadó leíró fogalomként vezették be az adatipar értékláncában elfoglalt helyükön változtatni kívánó ambiciózus piaci szereplők elemzői.¹ Olyan sikeresen talált aztán egymásra az üzleti és a technológiai jövőkép, hogy azóta is tematizálni tudja a médiaközbeszédet, s itt megerősödve immár szilárd terminus technicusként vonult be az információtechnológiai iparág (az *ICT-industry*) szakmai szótáraiba is.

A szociológia nem egyszerűen használni kezdte a frissen megnyíló adat-kisvilágokat, hanem azokat szisztematikusan beépítette saját módszertani fegyvertárába. S mivel ez az adaptáció egyszerre ment végbe a második generációs Nagy Adat-technológiák megerősödésével és az ezeken keresztül megnyíló újabb kutatási lehetőségek megjelenésével, megrázó erejű felismerésként érte a szociológusokat, hogy talán a társadalom (és a társadalmi cselekvés) megismerésének új korszaka érkezett el, amely – néhány dolgot legalábbis – gyökerestől változtat meg abban, ahogyan a „mesterséget” művelni kell. S bár örvendetesnek mondhatjuk, hogy az üzleti-technológiai interpretációk egyoldalúságát megtörve

¹ A tanulmány „az Európai Unió és Magyarország támogatásával, a TÁMOP 4.2.1.D-15/1/KONV-2015-0006 azonosító számú – Ösztöndíj magyar és külföldi hallgatóknak és kutatóknak – A közzei innovációs kutatóbázis és tudásközpont fejlesztése a Pannon Egyetem oktatási és kutatási hálózatának keretében” program részeként készült.

a társadalomelmélet számára is nagykorúsodott a fogalom,² úgy tűnik, hogy átugrottunk egy lépést. Már *azelőtt* *elkezdtük a Nagy Adatnak a saját diszciplinára gyakorolt hatáskövetkezményeit rendszerezni, mielőtt felfedeztük volna magunknak a Nagy Adatot magát, mint értelmezésre váró Jelenséget*, hogy abból önálló kutatási Tárgyat teremtsünk, belefeledkezve annak összetettségébe, szépségeibe és izgalmaiba. A Jelenséggel való szembesülésnek, a Tárgy fogalmi tisztázásának és beágyazásának elmaradása így programszerűen segítette az információtechnológia újabb fejleményeinek fogadtatásakor már megszokott gondolkodási és publicisztikai csapdák kialakulását. Ahogy már korábban annyiszor, kialakult a technooptimizmus-technopesszimizmus tengely, amelyen a Nagy Adat evangélistái látszanak szemben állni a veszélytudatosokkal. Utóbbiak ideológiának vagy vallásnak tekintik a „szép új adatvilág” prófétáinak lelkesedését, akikkel szemben örömmel tetszelegnek az orwelli forgatókönyvek feltáróinak, egyfajta kollektív lelkiismereti kapuőröknek a kisebbségi szerepében.³ Ha nem kezdené már az akadémiai közbeszédet is, önmagát kritikainak tartva, meghatározni ez a szemléleti kiindulópont és gondolkodói attitűd,⁴ s ha nem indulna ki egyre több publicista abból a feltételezésből, hogy a médiafigyelmet az apokaliptikus adatjövőre irányítva lehet felkelteni, akkor még gondolhatnánk azt, hogy időben vagyunk. De sajnos elkéstünk.

A következő oldalakon azt igyekszem kifejteni és indokolni, mit értek alatta és miért vélem úgy, hogy a társadalomtudományi jellegű diskurzus és a Nagy Adat közbeszédének fő-

2 Úgy is fogalmazhatnánk persze, hogy a társadalomelmélet „visszahódította” a kifejezést. Kevésbé ismert ugyanis, hogy a Nagy Adat használatának 2008-tól erősödő „felfutását” tíz évvel megelőzte John Mashey-nek, a Silicon Graphics vezető kutatójának (*chief scientist*) egy 1998-as előadása, ahonnan hódító útjára indult (Mashey 1998). Mindezt persze alapos nyomozó munkával sikerült kideríteni (Halevi és Moed 2012; Lohr 2013), mert a kézikönyvek sokáig az önjelölt Francis Dieboldhoz kötötték, aki makro-gazdaságtani munkáiban 2000-tól kezdődően fordult a terminushoz. Csakhogy a fogalom igazi megalkotója és legelső használója a szociológus, politológus és történész Charles Tilly (1929–2008), aki az írógép és a fénymásolat világába ájult társadalomkutatókat az egzotikusnak tűnő, mégis kézenfekvő gépi adatfeldolgozással elérhető állományok (Nagy Adat) és az azokhoz kapcsolódó nagyobb méretű kutatóközösségek (*bigger research tems*) növekvő fontosságára emlékezteti egy legendás kéziratában (Tilly 1980: 21). Tilly számos „legjobb gyakorlatot” fel is sorol, ami nem csoda, mert akkorra már a számítógép használata ugyancsak elterjedt volt bizonyos történettudományi műhelyekben. Egészen idáig csak a kifejezésről beszéltünk: maga a diskurzus már egy évtizeddel korábban, 1970-ben elindult egy tanulmánnyal (Halevi és Moed 2012: 3–6), csak még másképp nevezték a jelenséget: ez a szöveg egy 1969-ben végrehajtott légköri és óceáni szondázás eredményéről számolt be és fogalmazott meg szempontokat az adattömeg feldolgozásával kapcsolatos módszertani kihívás kezeléséhez.

3 A szarkazmus természetesen nem az adat-aktivizmusra (*data activism*), mint adekvát politikai reakcióra vonatkozik (Milan 2015). Az más kérdés ugyanakkor, hogy az adat-aktivizmus személyekhez fűződő egyedi sikereinek hitelességét erősen gyengíteni tudja, ha a nyilvános kommunikációban az egyértelműen azonosítható adattulajdonosi vagy adathasználói visszaélések konkrét elemzését általános ellenségképet konstruáló, kétes érvényességű, talmi szórványirodalomból összeállózott tételmondatok helyettesítik.

4 Dessewffy és Láng vitaindítója szerencsére mentes ettől az attitűdtől. Csalódást keltő volt átélni viszont, ahogy május végén, az információkutatók világkonferenciáján egy nagynevű és elismert pályatárs, Payal Arora (2015) a Stilgherrian néven publikáló ausztrál szabadúszó újságíró tudáshiányról árulkodó, csúsztatásoktól hemzsegő, mihaszna írásainak tételeit visszahangozta (Stilgherrian 2014a, 2014b). Ráadásul három tézise (hogy ti. az identitás, a demokrácia és a geográfia adatbázis-alapúvá (*database*) vált volna, erősen problematikus. Abból például, hogy a tradíció bizonyos aspektusai árúvá is váltak, még semmiképp nem következik, hogy szükségszerűen felülírja a személyes (identitás)választás mintázatait.

sodra mostanra *hermeneutizált, metonimizált és dekontextualizált* lett.⁵ Ennek nyomai még véletlenül sem fedezhetőek fel Dessewffy Tibor és Láng László remek és invenciózus vitaindító írásában, ám talán van jelentősége annak, hogy felhívjuk a figyelmet a kérdéskör néhány „logikai vakfoltjára”.

Hermeneutizálás: a jelenség megismerése helyett interpretációs játszmák

A Nagy Adat világa lényegét tekintve nem ideológia, nem vallás, nem jövőkép, nem gondolkodói narratíva: amennyiben valamilyen módon megfeleltethető volna is ezeknek, úgy csak másodlagosan. Merthogy elsősorban létező társadalmi gyakorlat: a diskurzus csak reflektálja azt, ahogyan egyre több alrendszer egyre változatosabban szorul rá a Nagy Adat megoldásaira. Az üzleti és társadalmi előrejelzések is végső soron pusztá hipotézisek, extrapolációk, forgatókönyvek: s az ezekkel szembeni ellenvélemények forrása is sokkal inkább a választott megközelítésmód mássága, és nem ugyanannak az „empíriának” az alternatív értelmezése. S még ahol a kiindulópont a reálszféra, ott is nagyon könnyű szinte észrevétlenül átcsúszni a „mi lenne, ha” (*what if*) világába. Ezt teszik példának okáért friss könyvükben Bernasek és Mongan (2015), akik a vállalatok azon törekvéseire hoznak példákat, amelyekkel azok mindenféle adatvarázslatokkal kipuhatólják, mi az az elérhető legmagasabb ár, amit a vevő hajlandó még megadni. A jól adatolt gazdasági szakmunkából azonban egyszer csak fikciós irodalom lesz, állatorvosi ló, amellyel jól illusztrálható, hogyan működik a hermeneutizálás.

Könyvük legfontosabbnak tartott (8.) fejezetében a szerzők egy képzeletbeli étterem képzeletbeli Nagy Adat-gyakorlatával kívánják szemléltetni, hogy az egykori törzshelyek személyes, emberléptékű világával szemben miféle elidegenedett rémálom vár majd a vendégekre. Ha már jártál ott, felismernek a külső kamerák, hogy odabenn már személyre szabott me-

5 Az ebből fakadó túlzások és torzítások korrekciója viszont önmagában most már kevés: túljutni akkor lehetne rajtuk, ha az alapkérdéseket *következetes és letisztult információs fogalmi* rendszerben, megfelelő *adatkultúra-történeti* beágyazással, *ismeretelméleti* erőterbe helyezve tesszük fel. Erre talán ennek a tanulmánynak a folytatásában nyílik majd mód. De addig is, hogy egyértelmű legyen, mire gondolok, álljon itt lapidáris stílusban néhány dolog, ami kiigazítást igényel. Inkonzisztens az „adat” fogalma, amióta a „rögzített információk egyik osztálya” jelentését a számítógépes adatfeldolgozás és a matematikai-statisztikai információelmélet, majd a melléjük gondolatresten felzárkózó könyvtár- és információtudomány agresszíven átalakította, és mindennemű, a feldolgozórendszerbe „belépő jel” értelmében kezdte használni. Evvel tovább növekedett az a fogalmi zűrzavar, ami az „információ” Shannon és Weaver-féle, kisajátításjellegű áthangolásával kezdődött meg, amikor az „új mentális mintázat” értelmű, minőségi természetű információból „zajos csatornán átvitt jel” értelmű, mennyiségi fogalom lett. Az egyszereggé tett DIK-modell (*data-information-knowledge*) alkalmatlan a sarkalatos összefüggések megértésére és szélsőségesen elavult, régóta megérett már egy alapos kritikára. A „mi, miért és mitől válhat adattá” kérdésre nem kielégítőek a válaszok, így az adatképzés helye sem rajzolódik megfelelően ki az információs viselkedés történetében, nincs „ontológiája”. Az eredeti, kultúrtörténeti beágyazottságú „adat” és a mögötte álló „adatkultúra” történeti feltárása csak részleges, így a teljes történeti ívre érvényes makromintázatok keresése is éppen csak megkezdődött (emiat is dicséretes, hogy Dessewffyék nem mulasztják el ennek a szempontnak a felvillantását sem). Ennek hiányában nem különül el például a memóriátámogatási és reprezentációtámogatási célú adattermelés, és nem látszik az óriási különbség sem az „emberi elmetartalom megszilárdult kimenete” és a „szenzorok révén az emberi érzékelést megkerülve feldolgozó rendszerekbe továbbított jel” között – ami például éppen a Nagy Adat kontextusában megkerülhetetlen jellegzetesség volna.

nüt kaphass tableten. Nem fogsz tudni ellenállni az első ital elfogyasztására kínált kedvezménynek, a többi meg majd jön magától. Az étterem mindentudó rendszerei tárolják majd, hogy mit ettél, mennyi ideig ettél, mit hagyta meg, a poharadról DNS-mintát rögzítenek, az okos beltéri kamerák még a széken való fészkelődés mintázataiból is hasznos adatokat generálnak, és előbb-utóbb az étteremnek gazdagabb képe lesz az egészségügyi állapotodról, mint az orvosodnak.⁶ E végső, legnagyobb, és az egész gondolatmenetet komolytalanra tevő túlzáson túlmenően némiképp megmosolyogtató az is, miért tartják valós lehetőségnek a szerzők a képzeletbeli étterem adatgyakorlatát, amely csak még nem létezik: mert azok a technológiák, amelyek az orwelli víziójukhoz kellene, már korai formában rendelkezésre állnak. De ez ugyebár már egy párhuzamos univerzum, nem a miénk. Hiszen ebben azt kéne végiggondolni, hogy vajon miért is akarna egy étterem működtetője elképesztő adatarzenállal és perszonalizációval (annak minden költségével együtt) mindenkiből egy kicsit még több fogyasztást kicsikarni, amikor ezt elég egyszerű módon is elérheti: finom ételekkel és jó áron kínált italokkal várva a betérőket. Arról nem is szólva, hogy az idősoros adatokhoz visszajáró vendég kell. A vendég meg akkor jár vissza, ha elégedett, ha az, amit kap, *value for money*.

De lépünk vissza az empiriához. Bernasekék hivatkoznak egy tanulmányra,⁷ amelyből megtudjuk például, hogy a Netflix pusztán a felhasználók alapvető demográfiai információinak felhasználásával 0,14%-kal növelte a profitját, a keresési előzmények felhasználásával pedig ez a szám 1% fölé is kúszhat. Most kéne ránézni a Valóságra: az online médiaszegmens más szereplőinél hogyan néz ki ez a dolog? Mi az ároptimalizálás felső határa? A minimális pluszköltség miként érinti a fogyasztói kosarakat? Milyen korrekciós mechanizmusok indulnak be? Hogyan viselkedik a konkurencia? Mi az értékmérleg: az online tartalomszegmens más területein elért jelképes árú vagy ingyenes fogyasztás ellensúlyozza-e az ebben a relációban tapasztalható minimális árnövekedést?

Az egyensúly „drámai eltolódását” vizionálni ehelyett az adatokkal gazdagon felvértezett eladók (*data-rich sellers*) felé még tűnhet pusztá túlzásnak, hevenyészett következtetésnek, a tisztességes áralku (*fair bargaining*) végét bejelenteni hatáskeltő nagyotmondásnak.⁸ Ám amikor kiderül, hogy a szerzők ezzel teszik kézzelfoghatóvá, hogy a „*vállalatok arra használ(hat)ják az adataidat, hogy kontrollálják a döntéseidet*”,⁹ akkor megérezzük, hogy

6 A Bernasek–Mongan-szerzőpáros több kisebb cikkben közölt részleteket a könyvből, ez az *A creepy vision of the new economic order, harvested from you* című darabból való (<http://www.pbs.org/newshour/making-sense/big-brother-watching-one-isnt-nsa/> letöltve: 2015 július 5.).

7 Ezt egy másik szatellit tanulmányban teszik (*Big data is coming for your purchase history – to charge you more money*), <http://www.theguardian.com/commentisfree/2015/may/29/big-data-purchase-history-charge-you-more-money>.

8 Az áralku részben sohasem volt „tisztességes” abban az értelemben, hogy a piacon kívüli eszközöket és az információs előnyt mindig is felhasználták a haszon maximalizálása érdekében (akárcsak a sokkal húsba vágóbb béralkuk során). Másrészt az árképzés sokkal összetettebb kérdés ma már, mint hogy a „mechanisztikus feltornázás” képzetével le lehetne írni. S végül a képlet nem is árról, hanem profitról szól: mi van (hogy Bernasekék kedvec fordulatával éljünk), ha a Nagy Adat technológiáinak egyik hatáskövetkezményeként olyan mértékben csökkenthetőek a termelési költségek, hogy az árak evvel nem arányos csökkentése okozza a profitnövekedést? És mi van, ha a Nagy Adat (vagy az ipari internet, az automatizálás vagy más) a fajlagosan egyre olcsóbb termékek irányába tolja az egész gazdaságot, amelynek minden olyan szegmense, amely pusztán ároptimalizálással kíván felszínen maradni, bukásra van ítélve? Vessük össze mindezt Rifkin (2014) bestsellerével, amely meggyőzően igazolja, hogy a költségek és az árak csökkenése a kortárs termelés és fogyasztás alapvető trendje.

9 Ez olyan erős tételmondat, hogy számtalan könyvajánló evvel a „lead”-del kezdődik. Lásd pl. a könyv Amazon-oldalát: <http://www.amazon.com/All-You-Can-Pay-Companies/dp/1568584741> (letöltve: 2015. július 8.).

nem a választott szcena jobb megismerésének vágya, hanem az antikorporatista panoptikum újabb viaszszoborral való gyarapítása a cél.¹⁰

Mindez egyáltalán nem jelenti azt, hogy a *Nagy Adat veszélydiskurzusa ne volna autentikus beszédmód*. Csak a Valóságot a maga összetettségében nem illusztrációnak kéne használni szedett-vedett „olvasatok” igazolásához. Székely Iván (2015) meg is mutatja hozzá szólásában, hogyan kell morális pániktól mentesen a valódi problémagócokra ráirányítani a figyelmet. Nagyon helytállóan érzem Stefania Milan téziseit is, aki együtt kezeli a Nagy Adat számos meglévő és kutatható sajátosságát (Milan 2015). Azt például, hogy miközben növeli a transzparenciát az egyik oldalon, gyengíti a participációs nyomást a másikon („ha úgyis itt a döntéseket támogató sok adat, nem kell az ember”). Hogy miközben növeli használói tudatosságát és hatékonyságát, információs egyenlőtlenségeket termel és erősít fel. Hogy miközben egyfajta egyetemes nyelvként (*universal language*) is funkcionál, nem az igazságot szállítja (*don't know the truth*), csak annak nyersanyagát.

S hasonlóképpen: az, hogy a Nagy Adatból hogyan lesz Nagy Profit, s hogy a vállalatok hányféleképpen tudnak előnyt kovácsolni a Nagy Adatból a maguk számára a növekedés vagy az innovációs potenciál erősítése érdekében, vizsgálható szakszerűen és tudományos módon is (Bögel 2011; Walker 2015). Mi sem bizonyítja ezt jobban, mint hogy Walker egyszerre ajánl keretrendszert, amellyel a Nagy Adatban rejlő lehetőségek jobban kiaknázzhatóak, és figyelmeztet azokra a jogi-szabályozási lépésekre, amelyek szükségesek ahhoz, hogy a féllégális megoldásokra hajlamos cégek kordában legyenek tarthatóak.

Metonimizálás: Egész helyett Rész

„*Dülöngél egy nagykabát*” Bródy János Koncz Zsuzsának írt dalában (Az első villamos), és mindeki tudja, hogy valójában a kabátba bújt emberre kell gondolnunk. Gyakorta szemléltetik ezzel a példával, hogyan veszi fel egy fogalom egy másiknak a jelentését a köztük lévő kapcsolatot felhasználva. A metonímia egyik alfaja a szinekdoché, ahol a névátvitel alapja a Rész-Egész viszony. Az irodalomból ismert metonimikus képtechika olykor a tudományban is hasznos, néha egyenesen heurisztikus tud lenni: egy jól megválasztott Rész segítségével könnyebben látótávolságba kerül az Egész. Ám mindez fordítva is „elsülhet”: ha az Egész helyett rosszul megválasztott Rész alapján teszünk állításokat, könnyűszerrel félrecsúszhat a diskurzus. Sajnos a Nagy Adat esetében ez számtalanszor megtörténik, nap mint nap, közleményről közleményre.

„A Nagy Adat nagy fenyegetést jelent az információ magánszférára nézve”, harsog a véleményvezérek egyik csoportja. Ha jó metonímiát használnának, akkor a Nagy Adat azokat az emberekben, ill. emberek csoportjaiban megtestesülő hatalmi és érdekkomplexumo(ka)t je-

10 Ha nem volna elég egyértelmű, mi a baj Hermészeinkkel: a döntések egyik, vásárlással kapcsolatos alosztályának adott eladó kínálatához kapcsoltnak jelentkező egyik döntéstípusának egyik összetevőjében tapasztalható kismértékű befolyásolhatóságnövekedés lehetőségéből univerzális (valamennyi eladót és valamennyi vevőt érintő) fogyasztói Armageddont növesztenek, ahol a fogyasztóból báb, a vállalatból bábjátékos lesz. Ki ne ismerne rá az elmúlt évtizedek kritikai reklám- és marketingkommunikációs irodalmának bornírt kiindulópontjára, amely kizárólag befolyásolható konzumidióták és ördögi manipulátorok egyoldalú játszmájának terepéként tudott a médiaszférára tekinteni? (Elfeledkezve arról, hogy az *egész kapcsolat az értéktérmezés és értékfogyasztás szférájába van lehorgonyozva*, amelyhez képest a reklámoknak a kommunikációs alrendszer egyik csatornájaként üzemelő kisvilága nem meghatározza és kontrollálja, hanem – jellemzően elhanyagolható, néha csekély, néha nagyobb mértékben – módosítja és befolyásolja a fogyasztói viselkedést.)

lentené, amelyek a polgárokkal szembeni ellenérdekek hordozói. Ám ezek a szerzők teljesen komolyan gondolják, hogy a fenyegetés forrása az Adat, így aztán nem a társadalmi konfliktus természetével, dinamikájával, csatornáival foglalkoznak, hanem mind hisztérikusabban irányítják a figyelmet sokkoló számokra és aggályos gyakorlatokra. Vagy itt van ez a másik végiggondolatlan metonímia: „a mesterséges intelligenciával párosított Nagy Adat átveszi a döntéseket az emberektől, dehumanizál, és így egy poszthumán jövő előfutára, megteremtője”. Szinte látjuk, hogy a Nagy Adat, mint Ágens – Tóth Árpád szavaival – „*testté tapad, alakká tornyosul*”, „*s Földet rengésre készítő hangja zeng: ím eljöttem*”. Igen ám, de ez megint csak rossz metonímia: az Adat nem cselekszik, nem hoz döntéseket. Azokat még a végletekig automatizált rendszerekben is emberek hozzák, a „ha-akkor” kapuk beállításával.¹¹ Az emberi figyelem és életidő felszabadítása a rutinszerű döntések alól, az adattechnika segítségével, egyesén humanizál.¹²

Vagy vegyünk egy unalomig ismert tételt: „a Google mindent tud rólad”. Ezt a fals metonímiát már annyit hallottuk, hogy meg sem rezzenünk, ha előkerül – pedig az egyik legpusztítóbb leegyszerűsítés.¹³ „*A Google tudja, mit gondolsz, mert elmenti a személyes kereséseidet*” (Schneier 2015).¹⁴ Jevgenyij Morozovtól mindez így hangzik: „A Google, a Facebook és az Apple (...) egyes kütyüi, alkalmazásai (...) azt is tudják, hova megyünk, mit olvasunk, mire keresünk rá, mi érdekel minket. *A Google tíz éve ül az adataimon: tudja, mit csináltam az elmúlt tíz évben, szinte azt is tudja, mit fogok a következő tízben.*” Steve Fuller pedig ezt teszi hozzá: „*Mark Zuckerberg talán jobban tudja, hogy ki vagy, mint te magad.*”¹⁵

11 „Az algoritmusokat emberek tervezik és építik, így azok készítöik egyoldalúságaira reflektálnak. Hogy egy régi szállóigét idézzünk: ha a dolgok mélyére akarunk nézni, végestelen végig újra és újra embereket látunk” (Gassée 2015).

12 Skót kutatók (Strathclyde Institute of Medical Devices) sebgyógyulást érzékelő vékony lapocskát fejlesztettek ki. Egy olyan szenzort, amelynek segítségével anélkül lehet információt nyerni a kötszer alatt amúgy „láthatatlan” seb állapotáról, hogy gyakori, szükségtelen és felesleges sérüléseket is okozni tudó átkötözésekkel kínoznák a beteget vagy gyarapítanak a kezelés költségét. Elég egy erre a célra kialakított monitorra pillantani, ami megjeleníti a fontos állapotértékeket, és jelez, amikor kötést kell cserélni. Ilyen és ennél bonyolultabb távdiagnosztikai rendszerek sora ellensúlyozza a figyelem szűköségét: hőmérsékletet, vérnyomást és más fontos indikátorokat mérő és az adatokat azonnal a feldolgozó központokba továbbító „automatizált” megoldások. Ha pedig a költségcsökkentést szem előtt tartó automatizációs düh olykor olyan területre téved, ahol feleslegesen és kockázatos módon helyettesíti az emberi döntéshozatalt a géppel, annak diszfunkcionális mivolta elég gyorsan kiderül. (A hatvanas évek közepe óta hosszú rekordja van mindennek.) Hasonlóképpen igaztalan dolog a Nagy Adaton számon kérni például a pénzüpiacok automatizált kereskedési rendszereit. Ott a tranzakciótípusokat lehetővé tévő szabályozás és a mögötte álló gazdaságfilozófia a „hunyo”. Még egy zsebmétonímia szokott kapcsolódni a fenti érvekhez: a Nagy Adat (illetve a ráépülő döntéshozatal) munkahelyeket szüntet meg. Ez azonban tévedés: az automatizációra igaz, hogy rövid távon munkahelyek esnek neki áldozatul, de a Nagy Adatra nem. A legfrissebb trend éppen az, hogy új típusú munkahelyek tömegét hozza létre, és éppen a mesterséges intelligencia korlátai miatt szükséges jelentéserzékenységi miatt (az új adatszakmákról ld. Z. Karvalics 2013b). Szemantikus technikák ide, automatikus hírszerkesztés oda, a nagy szöveges forgalmú webes oldalak például sorra állnak át (vissza) olyan munkatársak alkalmazására, akik „humán animátorként” magasabb felhasználói élményhez segítik a látogatókat.

13 Figyeljünk az alkímiára: hogyan lesz a *valamiből minden*, hogyan lesz egy rögzített „log”-ból tudás?

14 Ez már dupla metonímia: abból, hogy mit keresel, tudni lehet, hogy mit gondolsz? Kezdő humoristák ujjgyakorlatának is könnyű lenne ezer oldalról cáfolni. Ez olyan fokú ostobaság, hogy akármilyen könyv áll mögötte, olvasásra érdemtelenként kell elhajtanunk – így el sem juthatunk ahhoz a végképp komolytalan részhez, amely azt állítja, hogy a kormányok a megfigyelés eszközeit azért használják, hogy az egész világon veszélybe sodorhassák az embereket. („*Governments use surveillance to ... put people in danger worldwide.*”) Egy sokatmondó című bestsellerről beszélünk (Data and Goliath), nem volnék meglepve, ha már készülne a magyar fordítása (Schneier 2015).

15 A szövetszerű idézetek forrása a Morozov budapesti Brain Bar-előadásáról szóló összefoglaló: http://mandiner.hu/cikk/20150605_visszavehetjuk_a_hatalmat_a_technologiai_oriascegekto?utm_source=mandiner&utm_medium=link&utm_campaign=mandiner_201506 (letöltve: 2015. június 30.) Morozovot részletekbe menően bíráltam egy blogbejegyzésben (Z. Karvalics 2015), ebből néhány bekezdést csekély mértékben megváltoztatott formában vettem át.

Egyik idézett szerzőt sem menti fel, hogy csak visszasugározzák egy felületes, sok más tollforgató által is gerjesztett, végiggondolatlan és a rendszerszerűséget nélkülöző gondolatmenet pánikkeltő végeredményét. Kezdjük avval, hogy *a kütyük és az alkalmazások nem „tudnak” semmi mást, csak programozott feladatuknak megfelelően rögzíteni hálózati eseményeket.* Ezekhez olykor algoritmusok kapcsolódhatnak, amelyek bizonyos mintázatokhoz (mint korábban már láttuk is) „ha-akkor” szerkezetben rendelnek beavatkozásokat, evvel a felhasználók általában személyre szabott ajánlások vagy árazás formájában szembesülnek. Hogy a fogyasztó autonóm döntéseivel kapcsolatban mindez hordoz-e veszélyeket, milyeneket, vannak-e technikai evvel szemben, arról, mint korábban szintén láttuk, lehetne beszélni különböző gyakorlatok és ellengyakorlatok tapasztalatai alapján. Morozovék nem ezt teszik, hanem ránk erőltetik egy „alany” fenyegető képét, hol „rendszernek” nevezve, hol afféle „cégagyat” (*corporate mind*) vizionálva, „aki”, mint ellenérdekelt, ördögi intelligencia, ül valahol a távolban, figyel, és tudja, mit tesz, miért, és ebből mi következik majd a jövőbeni életedben, hogy ahhoz igazíthassa még ördögibb terveit. Csakhogy ami ezekben a rendszerekben történik, az személytelen és egyáltalán nem intelligens, nem tételez fel veled ellenérdekű tudatot. Nem „információs szuperhatalmakkal” állunk szemben, ahogy Székely Iván nevezi őket. Nem görnyed napestig monitorok felett szörnyű homunculusok armádiája, hogy egyenként vesézze ki online szokásai alapján a fogyasztókat: ehhez annyi homunculus kéne, ahány „megfigyelt”. Ezért kell csalni, túlozni, rossz metonímiákat használni, máskülönben nincs borzongató üzenet. Különösen az „általuk” tíz évre előre ismert jövő fenyegető víziója nevenséges: néhány múltbéli online tett (vásárlás vagy érdeklődő kattintás) alapján hogyan is lehetne előrejelzést tenni egy olyan fogyasztói viselkedésről, amelynek nagy része még nem létező vállalatok még nem ismert szolgáltatásainak vagy termékeinek teljességgel felmérhetetlen szabályok mentén történő használatáról szól, jócskán átalakuló kulturális és társadalmi környezetben? Zuckerbergnek meg a homunculusainak is számtalan jobb dolga van, mint hogy profilt alkosson Rólad, és jobban tudja, ki vagy.

De maga ez a „náluk vannak az adataid, szorongj”-mantra is felháborító leegyszerűsítés. *Online tranzakciós lábnymod nem Te vagy,* a magad gazdagságában és teljességében, a metonímia itt sem működik. Ez a lábnymom valójában személytelen jelsor, amelyet algoritmusok próbálnak, anonim és aggregált módon, más jelsorokkal együtt, értelmes mintázatokká formálni, amely az értékesítés komplex folyamatában olyan vállalati előnyre formálható, amely csak akkor működik, ha a fogyasztónak is személyes értéket ad. Ebben Te, Rettegésre Kondicionált Fogyasztó, személyiségednek csak töredékével vagy jelen, ahonnan bárki, bár hogyan indulna el visszafelé, fontosabb rétegekhez, pillanatok alatt eltévedne.¹⁶ A Google, a Facebook, az Apple és a többi cég nem személyesen érdeklődik utánad. Nincs nekik semmiféle erőforrásuk, hogy Veled foglalkozzanak. Számukra jóval hasznosabbak azok a *viselkedési makromintázatok,* amelyek birtokában finomítani tudják az algoritmusait, amelyekkel még jobb szolgáltatásokat tudnak nyújtani neked a még több profit reményében. Te pedig vagy hűséges vagy hozzájuk, és fizetsz, amíg jó neked, vagy elpártolsz tőlük, ha valami nem tetszik, és ha sokan tesznek így, akkor rossz neked.

16 Néhány napja az ebay-en keresztül, Ausztráliából rendeltem pár iPhone-os képernyővédő fóliát a legkisebb lányomnak, egy fülhallgatót meg a középsőnek. Ehhez a megrendeléshez az ékszer a legnagyobb, a cipő a fiamnak, és néhány filléres, de szép porcelán fogpiszkálótartó a gyűjteményembe, magamnak már Kínából indult útnak. Ha létezne a túloidalon homunculus, aki csendben a profilomat építené, most azt gondolnám rólam, hogy 45-ös láb méretű, metroszexuális zeneőrült vagyok, okostelefonok nagyfogyasztója, aki porcelán használati tárgyak vizsonteladásával foglalkozik. És még csak nem is vetette össze a Facebook-profilommal, ami amúgy nincs...

Talán úgy tűnhet, hogy a fenti példák túl egyszerűek, túl kézenfekvőek. Emiatt Csepeli György téziséhez – „a Nagy Adat a természeti és társadalmi folyamatok korábban elképzelhetetlen kiterjedésű és komplexitású modellezését teszi lehetővé” (Csepeli 2015: 171–176) – kapcsolódva mutatom be, hogyan siklatja ki az értelmes párbeszéd vonatát a rossz metonímia váltóállítása (különösen akkor, ha ahhoz még hermeneutizálás is társul).

Nézzünk egy kritikai állítást, amely látszatra antitezisként szolgál. „Széles körben elterjedt hiedelem, hogy a nagy adatállományok (*large data sets*) az intelligencia és a tudás magasabb formáit kínálják, amelyek olyan mélységű bepillantást eredményeznek (a természeti és társadalmi folyamatokba), ami korábban nem volt lehetséges” (Boyd és Crawford 2012: 663).

Az első gond abból adódik, hogy a megismerés új szintjére való lépés ígérete nem az általában vett adattömegekből fakad, nem „a” Nagy Adat-mivolt hordozza magában. Mindig konkrét valóságsgemzenekre vonatkozó konkrét „nagy adatok” eredményezik: nem az Egész, hanem a Rész. Induljunk el egy egzotikusnak tűnő példától. Elefántfókák fejére szerelt szenzorok és jeladók segítségével a közelmúltban 400 ezer adatprofil sikerült összegyűjteni az óceán vízhőmérsékletével és sűrűségével kapcsolatban (valós időben megosztva egy portálon). Ez egyike a számtalan hasonló projektnek: a Smithsonian például majmokkal, denevérekkel és méhekkel folytat szenzoros állatkövetés/adatgyűjtés kutatásokat (*animal tracking*), a Harvard tudósai massachusettsi fáknak adtak tömegesen „szemetet és füleket”. Korábban évtizedekig próbálkoztak hasonló megoldásokkal (madarakra szerelt rádiós adóvevővel például), de a jel előállítására és továbbítására egyidejűleg alkalmas és hosszabb időre elég energiabetáplálással rendelkező, kellőképp kicsi szerkezetek használata csak az utóbbi időben vált valóságos lehetőséggé (Kobie 2015). Az új tudás forrása egyáltalán nem a Nagy Adatban való hit (amit Stilgherrian, a hermeneutizálás hőse hübrisz hajtotta ideológiának becéz), hanem az új lehetőségekkel élve ténylegesen megtermelt adattömegek, amelyek révén új kisablakok nyílnak a Valóság megismerésére. Ez még jobban látszik akkor, ha a legnagyobb adattermelőkre, a tudományos óriásberendezésekre gondolunk, a genfi részecskegyorsítótól az elképesztő tömegű adatot produkáló űrtávcsőmonstrumokig a maguk exabyte-os kibocsátásával. Amíg ezek nem léteztek, az adataik révén megkomponálható ismeretek megszerzésére sem nyílt esély. S igaz mindez a mitokondriális DNS-vizsgálatokból nyert adatbázisokra, a világ hírtermésének sokváltozós követését lehetővé tévő adattárházakra, vagy akár a digitalizált, kereshető középkori képkönyvtárakra, az ötmillió digitalizált könyvben kereső Google Ngram Viewerre (Aiden és Michel 2013), de több ezer (!) hasonló természetű adatkisvilágra is.¹⁷

Metonimikus tévedés továbbá azt állítani, hogy a Nagy Adat, a Rész az új tudás forrása. A Nagy Adat előállításához szükséges technológia, a megtermelt adattömeg, annak tárolása és kezelése, a feldolgozás infrastruktúrája és az ennek kimenetét értelmező tudósközösség együtt alkotja azt a tudásinfrastruktúrát (Borgmann 2015), ami az Egésztest testesíti meg – az adatnak önmagában sem értéke, sem jelentése nincs.

Mivel az egyes tudományterületek adattermelési és felhasználási kultúrája rendkívüli módon különbözik, semmilyen okot és nyomást nem találunk, ami kiválthatna a Nagy Adat ál-

17 Arról most ne is essen szó, milyen sokatmondó, hogy a Nagy Adat dekonstruálójánál a Tudomány alrendszereiben működésbe állított különleges (Erik Brynjolfsson szavaival: briliáns) adattermelő kapacitás megismerő értéke megkérdőjeleződik, de a közösségi média felhasználói lábnyomaiból összegereblyézett banális adattömegekből „minden” megtudható.

talános fetisizmusát.¹⁸ Attól, hogy esetleg egy-egy felkészületlen publicista vagy diszciplináris határait feszegető kutató szavait lehet úgy értelmezni, hogy összegyúrható egy erre emlékeztető nézőpont,¹⁹ semmiképp nem ad útlevelet a metonimikus tigrisugráshoz, amellyel egy még véleményként is kisebbségi nézőponttal (a Rész Részével) kezdik el helyettesíteni, illetve kiszorítani a mindennapi tudományos gyakorlatban megnyilvánuló realitást (az Egészet).

S végül essen szó minden eddig említett metonímiagondok legnagyobbikáról is. A Nagy Adat valójában absztrakció, amely a kihívás és részben a használt megoldások közössége folytán az „általánost” testesíti meg. De már önmagában is rossz metonímia, mert a fogalmi innovációval megragadni kívánt Jelenségnek nem az adatállomány mérete az egyetlen attribútuma: ugyanolyan fontos a mérettel összekapcsolt kezelési-feldolgozási kihívás, ill. az abban rejlő újfajta megismerési lehetőség, az adatok termelődésének tempója, ill. megjelenítésük esetleges azonnalísága stb.²⁰

Emiatt a konkrét kifejtések mindig *valamilyen* Nagy Adat-világ (*dataverzum*²¹) sajátosságaira építenek, szükségszerűen Részre, miközben az állítások érvényessége igyekszik az Egész, a legmagasabb absztrakciós szint közelében maradni. Ezeket a dataverzumokat az előző bekezdésben felsorolt közös sajátosságok összekötik, de sokkal számosabbak azok a különbségek, amelyek elhasonítják őket egymástól. Aki nem fogalmaz elég körültekintően, s nem teszi egyértelművé, *melyik* dataverzumból beszél, az könnyűszerrel téved az általánosí-

18 Ennek pont az ellenkezője igaz: a tudósok tisztában vannak vele, hogy nem a Nagy, hanem a Megfelelő adat a fontos: ha ez éppen „kis adat” (*little data*), akkor az (Borgmann 2015). Emiatt az „öncélú adatszaporítás” vádjára épülő érvek sem helytállóak. Abból, hogy létezik olyan ismeretelméleti érv, amelyik azt mondja, hogy bármilyen rögzített adat válhat később újként felismert adatmintázat forrásává, és emiatt potenciálisan heurisztikus értéke lehet, még nem következik, hogy a tisztánlátásukat meg nem őrző tudósok félreteszik saját kutatási kérdéseiket, és minél haszontalanabb, ám annál nagyobb tömegű adatmeddőhányók (*data wastelands*) létrehozásában lelik örömeiket. Székely Iván (2015: 209–225) is túloz, amikor a „datafikáció korlátlan terjeszkedését” bírálja: néhány így értelmezhető megszólalásból nem következik az, hogy a valóságban is ez menne végbe. Kéjes örömmel lehet felmutatni egymástól távol fekvő valóságdarabkák számokra leképezhető mintázataiban fellelt azonosságok értelmetlenségét és hiábavalóságát, de méltatlan azt állítani, hogy valakik ezeket jelentéstartó és fontos produktumként kívánják letolni a védtelen torkunkon. (Arról nem is beszélve, hogy bizonyos adatvilágok egymás mellé sodródásának esetlegesen derűt keltő voltából nem következik az, hogy más, egymáshoz sodort adatvilágokban ne rejlene benne az új megismerő érték lehetősége.) Azon sem kell csodálkoznunk, hogy az egyik téves kiindulópont hozza a másikat: a mihaszna adattermelés vélt erődjére támadó keresztes lovagok sok időt töltenek el avval, hogy az adatbuborék (*data bubble*) kipukkanásának várható idejéről értekezzenek, a múlt évezred végén megélt ún. dotcom-lufi analógiájára. Vagy a „buta” Nagy Adat antiteziséként az „okos” adat (*smart data*) mellett kardoskodjanak (Rossi 2014).

19 A Nagy Adat fetisizmusának, amely a vélemények perifériáján helyezkedik el, nincs helye a komoly szakmai diskurzusban, nem véletlenül kapta a keresztségben a Pig Data (általánosan röfistatistikának fordított) nevet (Z. Karvalics 2013a). S mintha ugyanilyen szatirikus módon jelenne meg Pater Sparrow 1 című 2009-es filmjében is, amelynek forgatókönyve Stanislaw Lem *Az emberiség egy perce* című könyve alapján készült (Lem 1998 [1986]). A cselekmény egy könyv körül forog, amelyben a tudás fájának gyümölcséiként számok fejeznek ki és fejtenek meg mindent. Aki csak érintkezésbe kerül a könyvvel, megváltozik.

20 Emiatt hemzseg a Nagy Adat szakirodalma a versengő attribútumoktól, amelyeket általában az egy szóval megnevezett sajátosságok kezdőbetűiből komponálnak. Az attribútumlisták (amelyekben például a 3V-től már a 7V-ig jutottunk) sikeresen pontosítják a fogalom jelentését, amelynek valódi értelmét és dimenzióit elfedi a megnevezése. A Big Data egyszerű, világos és közérthető ugyan, erősen emlékeztet rokon terminusokra (*Big Science, Big Technology*), de kimondása láthatóan még nem elég a közös kódok tisztázásához, a párbeszédet lehetővé tévő közös jelentések kiformalódásához.

21 Az univerzum/mikroverzum szóparra emlékeztető módon képzett dataverzum (*dataverse*) szerencsés választásnak tűnik az egymástól független adatvilágok megnevezésére, még akkor is, ha a Harvard Egyetem nyílt forráskódú webalkalmazása tulajdonnévvé is tette. A *Dataverse* platformja egységes felületet kínál a tudományos eredmények megosztásához, megőrzéséhez, hivatkozásához és elemzéséhez. Ld. <https://dataverse.harvard.edu/>.

tás bozótosába. Vannak dataverzumok, amelyek nyíltak, és vannak, amelyek zártak. Vannak olyanok, amelyekhez kapcsolódik közvetlen gazdasági érdek, vannak olyanok, amelyekhez nem. Vannak olyanok, amelyek *privacy*-érzékenyek, mások ebből a szempontból semlegesek. Van, ahol a létrehozás, van, ahol a tárolás és visszakeresés, van, ahol a megjelenítés, és van, ahol az értelmezés/jelentésadás a kulcsmozzanat. Másképp működik a kísérleti adat (*experimental data*), mint a terepről gyűjtött (*field data*). Egészen más vonatkozik az ember által előállított, jelle kódolt adatra és a szenzor által közvetlenül feldolgozó rendszerekbe továbbított jelekre.

Sokkal több a független dataverzum, mint amennyi a szakmai-közéleti vitákban ténylegesen megjelenik vagy tárgyalatik. Méretüket tekintve a leggyakrabban tárgyalt megfigyelési-nyilvántartási és (online fogyasztói) lábnyom-másodfelhasználási kisvilágok a sor végén kullognak csak. Egy elképzelt lista élén rendre a Tudomány különböző adattermelő rendszerei tülekednek, majd nagy sokára megjelennek az ún. „*ipari internet*”-hez sorolható megoldások (a termelés és szolgáltatás világának operációit adatokkal követő és adatokra leképző rendszerei, a járműgyártók szervizadatbázisaitól a nemzetközi légi forgalom járat-, utas- és poggyászkövető rendszerein át a nemzetközi pénzügyi világ összekapcsolt adathálózatáig). Ha a weben elérhető digitalizált kulturális örökséget és a már digitálisan létrejövő tartalmak világát egységes állománynak tekintjük (*Corpus Digitale*), akkor ez fokozatosan zárkózik fel az előzőekhez. A „megosztott és/vagy cserélt tartalom” (saját fotók, bejegyzések, kommentárok, szövegek, zenék stb.) szintén webes dataverzuma (amelyhez elsősorban, de nem kizárólag a közösségi médiaként és kommunikációs célú alkalmazásként számon tartott felületek adattömege tartozik) a felhasználók számának monoton gyarapodásával nő. Más természetűek a keresési (*böngészési*) *adatrekordok*, megint mások az online *kereskedelem tranzakció-regisztráló* rendszerei. És akkor még ott vannak a különböző *intézményi adatgazdák*: multinacionális vállalatok, nagy szervezetek, kormányok, nagyvárosok, egyetemek, amelyek önálló adattavakba (*data lake*) gyűjtik – olykor előzetes strukturálás nélkül – az adataikat, hogy azokkal vagyonelemként (*data asset*) gazdálkodhassanak később (Woods 2011). Ha ezt a sokszínűséget (amit a választott technológiai és humánerőforrás-megoldások mássága csak tovább fokoz) nem vesszük figyelembe, szinte borítékolható a metonímiacsapdák kialakulása a diskurzusban.

Dekontextualizálás: a Lényeg távoltartása

A Nagy Adat jelensége önmagában értelmezhetetlen. Helye csak egy jóval átfogóbb mezőben, a *kortárs adatkultúra világán* belül jelölhető ki, értelme itt mutatkozik meg, magyarázatához innen kell szempontokat és érveket szállítani – emiatt az adat-ökoszisztémában a Nagy Adat mellett a Kis Adat, az Összekapcsolt Adat (*Linked Data*) és a Nyílt Adat (*Open Data*) együttes tárgyalása nélkülözhetetlen (Shiri 2014).²²

A kortárs adatkultúra maga pedig beágyazódik az *emberi civilizáció 21. század eleji kihívásairól*, adaptációs és innovációs kényszereiről és szükségszerűségeiről, a természeti és

²² A 2015 áprilisában rendezett ALLDATA-konferencia mintaszerű módon együtt tárgyalta ezt a négy különböző megközelítési módot (*The First International Conference on Big Data, Small Data, Linked Data and Open Data*; Barcelona, 2015. április 19–24.). Az adat-ökoszisztémáról, beleértve az adatvilág humán oldalának változásait is, részletesebben ld. Z. Karvalics (2013).

társadalmi környezetből leelkedő veszélyek elkerüléséről, fenntarthatóságról, a világrend diszfunkcióiról folytatott párbeszédbe és kollektív cselekvésekbe. Az adatvilágok mögött tehát az (elsődleges valóságként, husserli értelemben felfogott) életvilágok állnak, a maguk állapotaival és dinamikáival – Csepeli György szavaival: „a végtelenbe nyúlva” (2015: 173). A tét a folyamatok befolyásolása: új tudatosság alapján megváltoztatott koegzisztencia a természeti környezettel és a jelenlegi viszonyokat átrendező társadalmi innováció. Mindez egyidejűleg igényli a megismerő és a koordinációs erő fokozását – a természeti és társadalmi környezetben egyaránt. Három olyan „domaint” emelek ki, amelyek esetében a Nagy Adatok helyi értéke és funkciója, illetve az a kontextus, amelyen belül értelmezendők, jelentősen különböznek.

A Föld állapotváltozásának követése és komplexitásának feltárása többek között immár az emberiség kollektív erőfeszítéseként a Digitális Föld, majd a „következő generációs Digitális Föld” (*Next-Generation Digital Earth*) fantáziánév alatt fut (Craglia et al. 2008). Ez a Föld-megfigyelő rendszerek (legyenek azok légköri, oceanográfiai, földfelszíni, szeizmografikus, műholdas vagy más megoldások), a geoinformációs rendszerek, a GPS-adatok, a szenzorhálózatok, a különböző földrajzi alapú keresők integrációját jelenti. A kutatások jellegzetessége, hogy az ökológiai, klíma- és időjárás modellek nyersanyagául szolgáló változók megismeréséhez nélkülözhetetlen a hatalmas mennyiségű adat gyűjtése, amelyhez hatalmas számítási kapacitásra is szükség van. Nem csoda: a rendszerszint a bolygó, a túlélés szempontjából óriási jelentősége lehet minden előrelépésnek, amelynek révén befolyásolhatóak az egyensúlyi állapotok és az előrejelző, védekező, illetve öngyógyító képesség, valamint szétválaszthatóak a kozmikus, planetáris és antropogén hatások.

Ez utóbbi szempont miatt az extrém adatigényű csillagászati és fizikai alapú kutatások is ide csatolnak vissza. A Föld azonban egyúttal globális bio- és szociokulturális rendszer is, így az adatintegráció közvetlenül kiterjed a biomassa állapotára, valamint a népsűrűsége, migrációs mozgásokra vagy a szegénységeloszlásra is. A kihívás mélysége és összetettsége miatt az elsődleges kérdés az, miként lehet kiegészítő ismeretek esélyével kecsgetető új adatcsatornákat megnyitni, és hogyan lehet biztosítani azt a humán kapacitást, amely a növekvő adattömeg értelmezéséhez, feldolgozásához szükséges, lehetővé téve és előkészítve a modellbe illesztést és elméletalkotást.

Más dimenziók kerülnek előtérbe az ember mint biológiai lény kutatásához használt Nagy Adat-megoldások értékelésekor. A Humán Genom Projekt a világméretű, erőforrás-megosztáson és tervezésen alapuló programok egyik úttörője volt, egyfajta „legjobb Nagy Adatgyakorlat”, s a folytatódó, ráépülő genetikai kutatások párhuzamosan ígérnek az öröklődés alapfolyamatainak jobb megismerését és a személyre szabott génterápiák új világát. Ennek a személyre szabottságnak, az egyéni egészségállapot folytatólagos monitoringjának az egyéni életesélyek és életminőség javítása a tétje.²³ Közép- és hosszú távon azonban az idősoros betegadatok egyre értékesebb általános diagnosztikus óriás adatbázisban halmozódnak fel, és új generációs szakértői rendszerek kialakítását segíthetik. Az új ismeretek megszerzésének azonban az orvosbiológiai területen a Nagy Adaton kívül számos más formája is van, és a Nagy Adat használt eljárásai is sok esetben a már felhalmozott tudáshoz való gyors hozzáférést segítik.

23 „Mindössze” arra kell figyelni, hogy más adatkörökkel (például biztosítás vagy adózás) ne kerülhessenek fedésbe az egészségügyi adatok.

Úgy vélem végül, hogy napjaink *alapvető társadalmi-gazdasági problémáinak* nagyobb felbontású leírása és megértése *nem a Nagy Adat technikáitól* remélhető, még annak Dessewffyk által erős programnak nevezett válfajától sem. S noha, ahogy Csepeli György írja, a globalizáció és az internet, mint szuperplatform minden szociológiai probléma kontextusát megváltoztatta, s „cseppfolyóssá váltak a határok” (2015: 172), a magas rendszer-szintekhez kapcsolódó modell- és elméletalkotásban a hozzáadott érték forrását a hagyomány ismeretéből építkező fogalmi és módszertani innováció, a kreativitás és a jól tesztelhető új hipotézisek környékén továbbra is lehet és kell is keresni. És még a Dessewffyk által „szociológia hagyományos nagy kérdéseiként” (a társadalmi egyenlőtlenségek, a kirekesztés, a hatalom vagy a társadalmi tőke vizsgálata) kiemelt területek esetében is inkább kiegészítő-támogató, mint új utakat nyitó szerepe lehet a Nagy Adat technikáinak (Dessewffyk és Láng 2015: 168). Sőt; ahol az eredmények normatívak, és azok szociálpolitikai vagy aktivista jellegű lefordítása után a folyamatokat befolyásoló beavatkozások feltételeinek megteremtése a kérdés, ott sem új adatkörök bevonása, hanem a meglévő és a valóság változásaival párhuzamosan, narratív és diszkurzív módon is gyarapodó ismeretek újszerű felhasználása eredményezhet tudástöbbletet.

Eközben kétségtelen, hogy *mindenhol, ahol az online viselkedés követhetővé és adatolhatóvá vált, az empirikus szociológia hatalmas lehetőségekhez jutott*, hogy tartalmak előállításán, fogyasztásán, egyes tranzakciótípusokon, geolokációs azonosításokon keresztül részben a régi módszereket leváltva, részben új utakat járva bővítse az arzenálját. Egy fokkal óvatosabb lennének azonban még a Nagy Adatnak evvel a „gyenge programjával” is, mint Dessewffyk – „a digitális lábnyomokból nyilvánvalóan kirajzolódnak fő csapásirányok” (2015: 168) – vagy mint Csepeli György – „az internetre került társadalmi folyamatok minden történése nyomot hagy maga után” (2015: 173) –, mert azért ismét a Rész Részéről beszélünk. A társadalmi folyamatok közül ugyan kétségkívül egyre több kerül az internetre, de nagyon sok nem. Ami az internetre kerül, annak csak bizonyos aspektusai tanulmányozhatók Nagy Adatokkal. És a Nagy Adatok révén tanulmányozható viselkedéseknek és sajátosságoknak vannak olyan vonatkozásai, amelyekhez remek forrás a Nagy Adat, de vannak olyanok, amelyek esetében irreleváns (például a pusztán nyomhagyásból nem következik a nyomhagyást eredményező szituáció rekonstruálhatósága). Nem feltétlenül kapunk választ az „okok” kérdésére, ha következményt látunk. A viselkedések azonosságára tudhatunk magyarázatot adni, de hogyan jutunk közelebb a különbségekhez? Esetleg megtalálható és felmutatható lesz *néhány eltérés jól klaszterezhető csoportok esetében, de hogyan mondunk érvényeset az egyéni viselkedések különbözőségére?*

Mint látható, nagyon sok a nyitott kérdés – de legalább van már narratívánk, amely értelmes és izgalmas párbeszédnek nyithat utat.

Hivatkozott irodalom

- Aiden, Jerez és Jean-Baptiste Michel (2013): *Uncharted. Big Data as a Lens on Human Culture*. New York: Riverhead Books.
- Arora, Payal (2015): *Big Data Commons and the Global South*. (Keynote előadás: The Information Society at the Crossroads, Bécs, 2015. június 6.)
- Bernasek, Anna és D. T. Mongan (2015): *All You Can Pay. How Companies Use Our Data to Empty Our Wallets*. New York: Nation Books.

- Borgmann, Christine L. (2015): *Big Data, Little Data, No Data. Scholarship in the Networked World*. Cambridge, Mass.: The MIT Press.
- boyd, danah és Kate Crawford (2012): Critical Questions for Big Data. *Information, Communication & Society* 15(5): 662–679.
- Bógel György (2011): Az adatrobbanás mint közgazdasági jelenség. *Közgazdasági Szemle* 58(10): 877–889.
- Craglia, Massimo et al. (2008): Next-Generation Digital Earth. *International Journal of Spatial Data Infrastructures Research* (3): 146–167. Interneten: <http://ijdsdir.jrc.ec.europa.eu/index.php/ijdsdir/article/viewFile/119/99>.
- Csepeli György (2015): A szociológia és a Big Data. *Replika* (92–93): 171–176.
- Dessewffy Tibor és Láng László (2015): Big Data és a társadalomtudományok véletlen találkozása a műtőasztalon. *Replika* (92–93): 157–170.
- Gassée Jean-Louis (2015): Human Curation Is Back. *Monday Note* (2015. július 5.). Interneten: <http://www.mondaynote.com/2015/07/05/human-curation-is-back/>.
- Halevi, Gali és Henk F. Moed (2012): The Evolution of Big Data as a Research and Scientific Topic. Overview of the Literature. *Research Trends* (30): 3–6. Interneten: http://www.researchtrends.com/wp-content/uploads/2012/09/Research_Trends_Issue30.pdf.
- Kobie, Nicole (2015): The Internet of Seals. How Sensors for Elephant Seals Tackle Climate Change. *The Guardian* (2015. június 17.). Interneten: <http://www.theguardian.com/technology/2015/jun/17/elephant-seals-sensors-climate-change-internet-of-things>.
- Lem, Stanislaw (1998 [1986]): *Az emberiség egy perce*. Budapest: Európa.
- Lohr, Steve (2013): The Origins of 'Big Data'. An Etymological Detective Story. *The New York Times Bits Blog*. Interneten: <http://bits.blogs.nytimes.com/2013/02/01/the-origins-of-big-data-an-etymological-detective-story/> (letöltve: 2015. július 1.).
- Mashey, John R. (1998): *Big Data... and the Next Wave of Infrastrass*. Interneten: https://www.usenix.org/legacy/event/usenix99/invited_talks/mashey.pdf (letöltve: 2015. július 5.).
- Milan, Stefania (2015): *Big Data and the Understanding of the Political*. In *the Session: How Smart is Big Data?* (Konferencia-előadás: The Information Society at the Crossroad, Bécs, 2015. június 6.)
- Rifkin, Jeremy (2014): *The Zero Marginal Cost Society. The Internet of Things, Collaborative Commons, and the Eclipse of Capitalism*. New York: Palgrave Macmillan.
- Rossi, Ben (2014): Is Big Data Dead? The Rise of Smart Data. *Information Age* (2014. szeptember 23.). Interneten: <http://www.information-age.com/technology/information-management/123458486/big-data-dead-rise-smart-data>.
- Schneier, Bruce (2015): *Data and Goliath. The Hidden Battles to Collect Your Data and Control Your World*. New York: W. W. Norton & Company.
- Shiri, Ali (2014): Linked Data Meets Big Data. A Knowledge Organization Systems Perspective. *Advances In Classification Research Online* 24(1): 16–20. Interneten: https://www.researchgate.net/publication/259889208_Big_Data_Meets_Linked_Data_A_Knowledge_Organization_Systems_Perspective.
- Stilgherrian (2014b): Why Big Data Evangelists Need to be Reprogrammed. *ZDNet* (2014. szeptember 30.). Interneten: <http://www.zdnet.com/article/why-big-data-evangelists-need-to-be-reprogrammed/> (letöltve: 2015. július 3.).
- Stilgherrian (2014a): Big Data is Just a Big, Distracting Bubble, Soon to Burst. *ZDNet* (2014. július 11.). Interneten: <http://www.zdnet.com/article/big-data-is-just-a-big-distracting-bubble-soon-to-burst/> (letöltve: 2015. július 3.).
- Székely Iván (2015): Az adatmentes zónák szükségessége és esélye. Helytelen reflexió Dessewffy Tibor és Láng László írására. *Replika* (92–93): 209–225.
- Tilly, Charles (1980): *The Old New Social History and the New Old Social History*. (Munkaanyag.) Ann Arbor: University of Michigan, 1980 October <http://deepblue.lib.umich.edu/bitstream/handle/2027.42/50992/218.pdf>.
- Walker, Russell (2015): *From Big Data to Big Profits*. Oxford: Oxford University Press.
- Woods, Dan (2011): Big Data Requires a Big, New Architecture. *Forbes* (2011. július 7.). Interneten: <http://www.forbes.com/sites/ciocentral/2011/07/21/big-data-requires-a-big-new-architecture/>.
- Z. Karvalics László (2013a): Pig Data – egy kis adatmalackodás. *IT Business Online* (2013. június 6.). Interneten: http://www.itbusiness.hu/Fooldal/publicisztika/Z_Karvalics_Laszlo/Pig_Data-egy_kis_adatmalackodas.html.
- Z. Karvalics László (2013b): Data Hugging – a magunkhoz szorított adatokon innen és túl. *Tudományos és Műszaki Tájékoztató* 60(7): 300–303. Interneten: http://tmt.omikk.bme.hu/show_news.html?id=5804&issue_id=550.
- Z. Karvalics László (2015): Anti-Morozov, avagy ami történik az Interneten, és ami nem. *IT Business Online* (2015. június 8.). Interneten: http://www.itbusiness.hu/Fooldal/publicisztika/Z_Karvalics_Laszlo/20150608_ZKL_Morozov.html.